

Introducing WAISserver™ 2.0

Power Tools for Professional Online Publishers

WAISserver™ is the leading publishing system in the Internet industry. Using WAISserver, organizations such as Dow Jones, the Library of Congress, Novell, CMP Publications, and Encyclopaedia Britannica have given Internet users access to books, magazines, news information, product data sheets, technical overviews, company information, and more. WAISserver allows content providers to index and publish large databases to the Internet, where they can be searched from around the world. Content providers can publish databases in remote locations and different formats, documents in different formats, and manage updates to databases without re-indexing.

Internet users can access WAIS databases with popular clients such as NCSA Mosaic, Netscape, and Gopher, and most importantly, search for the information they need by typing in simple natural language queries.

HISTORY OF WAIS

WAIS (Wide Area Information Servers) is the standard for Internet publishing. WAIS Inc. founder Brewster Kahle developed WAIS as a freeware product. Two years ago he formed WAIS Inc. to redesign WAIS for large scale publish-

ing. The commercial version, called WAISserver, goes far beyond the capabilities of the freeware version, which helped establish WAIS as a standard for Internet-based online publishing. There are now over 800 WAIS databases on the Internet today.

WAIS AND THE WORLD WIDE WEB

WAIS is the key that unlocks the content behind the World Wide Web. WAIS provides searching capabilities to point-and-click clients like Mosaic, in which users search for information by navigating hypertext links between documents. For instance, suppose a company wanted to buy a copy machine with specific features which was small enough to fit in a utility closet.

An employee of the company could go to a vendor's home page on the Web, and if it is "powered by WAISserver", type in "Tell me about copiers with two sided copying, and automatic collating, no bigger than 48 inches by 36 inches," get all the available models, and ask for pricing and distributors. With a Web server that wasn't powered by a WAISserver, the end user would have to navigate links and hope to find the relevant information.

What's WAIS About This Picture?

WAISserver Components

WAISserver 2.0 includes eight major functionality modules with each license:

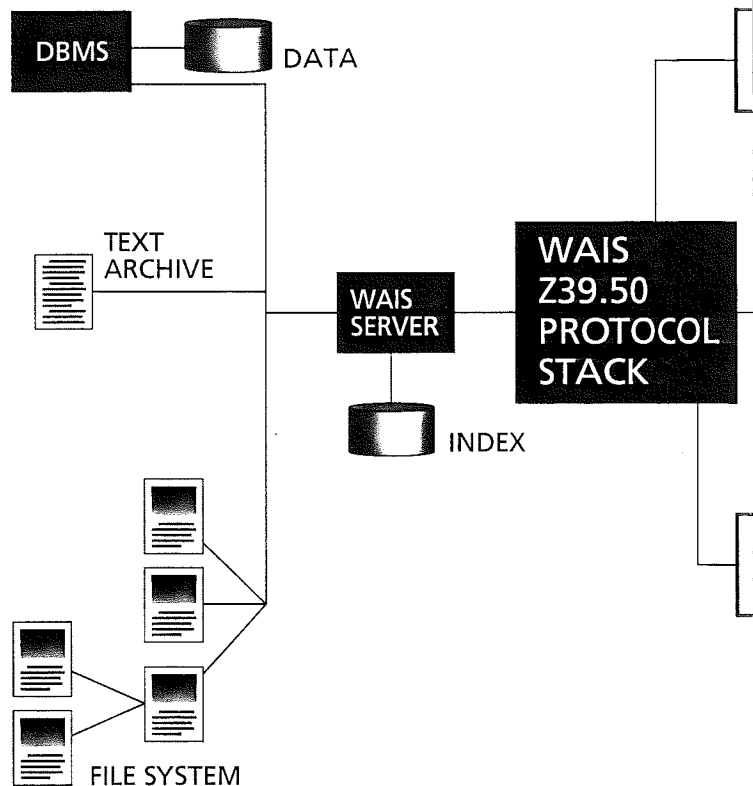
- **CUSTOM PARSER TOOLKIT** - a customizable API for getting content from documents in any type of format.
- **INDEXER** - for indexing local and remote content databases.
- **SEARCH AND RETRIEVAL SOFTWARE** - search functionality for local and remote databases, with relevance ranking and relevance feedback features.
- **THE WAIS INC. Z39.50 PROTOCOL STACK** - includes compatibility with versions one and two.
- **WAISgate** - a gateway that converts the HTTP protocol used by the World Wide Web to Z39.50 so that World Wide Web (WWW) servers and clients (such as NCSA Mosaic, Gopher, and Netscape) can initiate a WAIS search.
- **GOPHER GATEWAY** - a gateway that converts Gopher protocol to Z39.50 so that Gopher servers and clients can initiate a WAIS search.
- **FORWARDER** - allows clients and servers separated by firewalls to access one another.
- **MANAGEMENT REPORTING** - reports containing more than 20 statistics pertaining to users, usage patterns, connect time, searches performed, and more.

Where WAISserver Fits in the Internet

The WAISserver software architecture has four main components: the client, the server, the database, and the protocol, as shown below. The client is a user-interface program that sends search and retrieval requests to local or remote servers. Most popular clients are supported, such as NCSA Mosaic, Netscape, and Gopher. The WAISserver product also includes a client.

The server is a program that services client requests. Servers are available for a variety of UNIX platforms. The server generally runs on a machine containing one or more information sources, or WAIS databases.

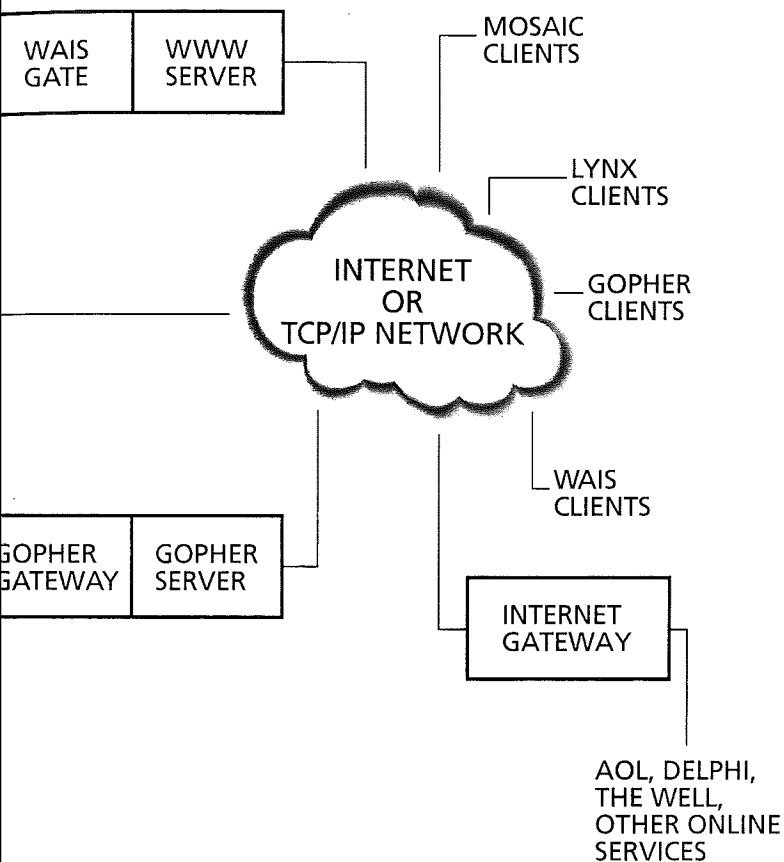
Where The W



When a client initiates a search request, the server searches an index rather than the databases where information is physically stored. WAISserver indexes databases automatically for a variety of different kinds of searches.

The WAIS protocol is used to connect WAIS clients and servers. It is based on the NISO Z39.50 Version 2 Information Retrieval Service and Protocol Standard. Z39.50 protocol allows the server to "speak" to databases that use other protocols, such as Web and Gopher databases via gateways.

WAISserver Fits



- **Date Range Searching**
Search for date ranges for precise control over timely information.
- **Delete Document**
Delete documents from a WAISserver database without re-indexing, for active Internet services.

LARGE DATABASE FEATURES

- **Multiple Databases**
Databases can span multiple disks.
- **Database Integration**
Access documents stored in traditional RDMS databases such as Oracle without re-indexing or re-architecting your system, allowing you to publish on the Internet with a minimum of engineering intervention.

CONNECTIVITY FEATURES

- **z39.50 Version 2 Support**
WAISserver supports the latest version of the z39.50 protocol, enabling your WAISserver to "speak" the standard search client languages.
- **WAISsearch Client (waislookup)**
New command line client for system administrators.
- **WAISgate**
Allows World Wide Web clients to access WAIS databases.

SEARCH FEATURES

- **Literal Search**
Limits search results to entries that match search parameters exactly.
- **Boolean Search**
And/or criteria.
- **Relevance Weighting**
WAISserver ranks documents in order of relevance to the query. Relevance weighting takes into account advanced indexing information such as key word proximity, density (number of times the keywords appear, and where), and locations of the keywords in the documents.
- **Relevance Feedback**
Can refine a search by indicating a document that closely matches the desired search results, such as "get me more documents like this one."
- **Query Report**
Lists terms searched, how the server interpreted the request, how many words and documents are being searched, and how many times each word appeared, and more.

WAISserver Features

PRODUCTION ON-LINE PUBLISHING FEATURES

- **Large Databases**
Index is limited only by the maximum size of available disk space.
- **Logging Capabilities**
Sophisticated reporting on usage statistics per database.
- **Custom Parser Toolkit**
Customize the indexing of any type of document.
- **Small Index Size**
One third the size of the data.

ACTIVE SERVICE FEATURES

- **Incremental Indexing**
Add new information to the database without re-indexing the entire database, for Internet publishers who have data that feeds in frequently.

WAISserver Benefits

BENEFITS TO INTERNET PUBLISHERS

WAISserver allows the Internet publisher to:

- Publish information to the Web without costly HTML coding
- Access local and remote databases
- Access large databases
- Automatically index databases
- Index multiple databases, and in effect "combine" them
- Mix proprietary databases to leverage existing MIS architecture
- Add new content without re-indexing the entire database
- Create HTML documents "on the fly"
- Publish documents in any native format through the use of an API
- Support a variety of popular clients
- Minimize start up and maintenance costs for an Internet-based service
- Track usage and status

BENEFITS TO INTERNET USERS

The WAISserver allows Internet users to:

- Search Internet databases by simply typing questions
- Perform literal searches, Boolean searches, and natural language searches
- Get the most relevant documents, listed in order of relevance
- Get query report on search requests, such as number of times a word appeared, and where
- Receive documents that satisfy search requests with their criteria highlighted in bold



690 Fifth Street, San Francisco, CA 94107 Web URL: <http://www.wais.com/> 415-356-5400 info@wais.com

WAISserver is a registered trademark of WAIS, Inc. All registered and unregistered trademarks are property of their respective owners. 12/94 v2.0